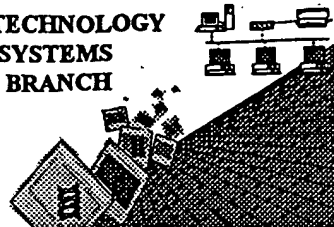


RAW SEQUENCE LISTING ERROR REPORT

BIOTECHNOLOGY
SYSTEMS
BRANCH



#. 12/1008
4/12/02

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/913,858

Source: 1600

Date Processed by STIC: 4/3/02

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4416.

PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER
VERSION 3.1 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND
TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<http://www.uspto.gov/ebc/efs/downloads/documents.htm>), EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202
3. Hand Carry directly to:
U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7th Floor, Examiner Name, Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202
Or
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 01/29/2002

RECEIVED

APR 11 2002

TECH CENTER 1600/2900

RECEIVED

APR 11 2002

TECH CENTER 1600/2900

Raw Sequence Listing Error Summary

ERROR DETECTED

SUGGESTED CORRECTION

SERIAL NUMBER: 09/913,858

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1 Wrapped Nucleics
 Wrapped Aminos The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
- 2 Invalid Line Length The rules require that a line not exceed 72 characters in length. This includes white spaces.
- 3 Misaligned Amino
 Numbering The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
- 4 Non-ASCII The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
- 5 Variable Length Sequence(s) contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
- 6 PatentIn 2.0
 "bug" A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
- 7 Skipped Sequences
 (OLD RULES) Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence:
 (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
 (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
 (ii) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
 This sequence is intentionally skipped

 Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
- 8 Skipped Sequences
 (NEW RULES) Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence.
 <210> sequence id number
 <400> sequence id number
 000
- 9 Use of n's or Xaa's
 (NEW RULES) Use of n's and/or Xaa's have been detected in the Sequence Listing.
 Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
 In <220> to <223> section, please explain location of n or Xaa; and which residue n or Xaa represents.
- 10 Invalid <213>
 Response Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence
- 11 ✓ Use of <220> Sequence(s) 3 missing the <220> "Feature" and associated numeric identifiers and responses.
 Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.
 (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
- 12 PatentIn 2.0
 "bug" Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
- 13 Misuse of n n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.

AMC/MH - Biotechnology Systems Branch - 08/21/2001

APR 11 2002

TECH CENTER 1600/2900

RECEIVED

APR 11 2002

TECH CENTER 1600/2900

RECEIVED



Does Not Comply
Corrected Diskette Needed

1600

Error on pg. 3

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/913,858

DATE: 04/03/2002

TIME: 10:22:05

Input Set : A:\030560-057.ST25.txt

Output Set: N:\CRF3\04032002\I913858.raw

RECEIVED
APR 11 2002
TECH CENTER 1001220

```

4 <110> APPLICANT: Altmann, Friedrich
6 <120> TITLE OF INVENTION: Fucosyl Transferase Gene
8 <130> FILE REFERENCE: 030560-057
10 <140> CURRENT APPLICATION NUMBER: US 09/913,858
11 <141> CURRENT FILING DATE: 2001-08-20
13 <150> PRIOR APPLICATION NUMBER: PCT/AT00/00040
14 <151> PRIOR FILING DATE: 2000-02-17
16 <150> PRIOR APPLICATION NUMBER: AT A 270/99
17 <151> PRIOR FILING DATE: 1999-02-18
19 <160> NUMBER OF SEQ ID NOS: 17
21 <170> SOFTWARE: PatentIn version 3.1
23 <210> SEQ ID NO: 1
25 <211> LENGTH: 2198
26 <212> TYPE: DNA
27 <213> ORGANISM: Unknown Organism
29 <220> FEATURE:
30 <223> OTHER INFORMATION: Description of Unknown Organism: plant
32 <400> SEQUENCE: 1
33 actaactcaa acgctgcatt ttotTTTTTtC tttcagggaa ccatccaccc ataacaacaa 60
34 aaaaaaacaac agcaagctgt gTTTTTTTTtA tcgttctttt tcttttaaaca agcacccecca 120
35 tcatggaatc gtgctcataa cgccaaaatt ttccatttcc ctttgatttt tagttttattt 180
36 tgcggaattg gcagttgggg gcgcaattga atgatgggtc tgttgacgaa tcttcgagggc 240
37 tcgagaacag atggtgcccc acaagacagc ttaccggtt tggctccggg aggcaaccca 300
38 aagaggaaat ggagcaatct aatgcctctt gttgttgccc ttgtggtcat cgcggagatc 360
39 gcgtttcttg gtaggttgga tatggccaaa aacgccgcca tgggtgactc cctcgctgac 420
40 ttcttctacc gctctcgagc ggtcggtgaa ggtgacgatt tggggttggg tttggtggct 480
41 tctgatcgga attctgaatc gtatagttgt gaggaatggt tggagagga ggatgctgtc 540
42 acgtattcga ggggcttttc caaagagcct atttttgtt ctggagctga tcaggagtgg 600
43 aagtcgtgtt cggttggtatg taaatttggt tttagtggg atagaaagcc agatgccgca 660
44 tttgggttac ctcaaccaag tggaacagct agcattctgc gatcaatgga atcagcagaa 720
45 tactatgctg agaacaatat tgccatggca agacggagg gatataacat cgtaatgaca 780
46 accagtctat ctteggatgt tctgttgga tatttttcat gggctgagta tgatatgatg 840
47 gcaccagtgc agccgaaaac tgaagctgct cttgcagctg ctttcatttc caattgtggt 900
48 gctcgaaatt tccggttgca agctcttgag gcccttgaaa aatcaaacat caaaattgat 960
49 tcttatggtg gttgtcacag gaaccgtgat ggaagagtga acaaagtga agccctgaag 1020
50 cactacaaat ttagcttagc gtttgaaaat tcgaatgagg aagattatgt aactgaaaaa 1080
51 ttcttccaat cccttggtgc tggaactgtc cctgtggttg ttggtgctcc aaatattcag 1140
52 gactttgtct cttctcctgg ttcaatttta catattaaag agatagagga tgttgagtct 1200
53 gttgcaaaga ccatgagata tctagcagaa aatcccgaag catataatca atcattgagg 1260
54 tggaagtatg agggctccatc tgactccttc aaggcccttg tggatatggc agctgtgcat 1320
55 tcatcgtgcc gtctttgcat tcaattggcc acagtgagta gagagaagga agaaaataat 1380
56 ccaagcctta agagacgtcc ttgcaagtgc actagagggc cagaaaccgt atatcatatc 1440
57 tatgtcagag aaaggggaag gtttgagatg gagtccattt acctgaggtc tagcaattta 1500

```

RAW SEQUENCE LISTING

DATE: 04/03/2002

PATENT APPLICATION: US/09/913,858

TIME: 10:22:05

Input Set : A:\030560-057.ST25.txt

Output Set: N:\CRF3\04032002\I913858.raw

```

58 actctgaatg ctgtgaaggc tgctgttggt ttgaagttca catccctgaa tcttgtgcct 1560
59 gtatggaaga ctgaaaggcc tgaagttata agagggggga gtgctttaaa actctacaaa 1620
60 atatacccaa ttggcttgac acagagacaa gctctttata ccttcagctt caaaggtgat 1680
61 gctgatttca ggaagtcactt ggagaacaat ccttgtgcca agtttgaagt catttttgtg 1740
62 tagcatgctc taaatggtac ctctgctcta cctgaattag cttcacttag ctgagcacta 1800
63 gctagagttt taggaatgag tatggcagtg aatatggcat ggctttatatt atgcctagtt 1860
64 tcttggccaa ctcatgatg ttttgtataa gacatcacac tttaatttta aacttgtttc 1920
65 tgtagaagtg caaatccata tttaatgctt agtttttagtg ctcttatctg atcatctaga 1980
66 agtcacagtt cttgtatatt gtgagtgaag actgaaatct aatagaagga tcagatgttt 2040
67 cactcaagac acattattac ttcatgttgt tttgatgac tcgagctttt ttagtgtctg 2100
68 gaactgtccc tgtggtttga gcacctgtta ttgcttcagt gttactgtcc agtggttatc 2160
69 gtttttgacc tctaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2198
71 <210> SEQ ID NO: 2
72 <211> LENGTH: 510
73 <212> TYPE: PRT
74 <213> ORGANISM: Unknown Organism
76 <220> FEATURE:
77 <223> OTHER INFORMATION: Description of Unknown Organism:plant
79 <400> SEQUENCE: 2
80 Met Met Gly Leu Leu Thr Asn Leu Arg Gly Ser Arg Thr Asp Gly Ala
81 1 5 10 15
83 Gln Gln Asp Ser Leu Pro Val Leu Ala Pro Gly Gly Asn Pro Lys Arg
84 20 25 30
86 Lys Trp Ser Asn Leu Met Pro Leu Val Val Ala Leu Val Val Ile Ala
87 35 40 45
89 Glu Ile Ala Phe Leu Gly Arg Leu Asp Met Ala Lys Asn Ala Ala Met
90 50 55 60
92 Val Asp Ser Leu Ala Asp Phe Phe Tyr Arg Ser Arg Ala Val Val Glu
93 65 70 75 80
95 Gly Asp Asp Leu Gly Leu Gly Leu Val Ala Ser Asp Arg Asn Ser Glu
96 85 90 95
98 Ser Tyr Ser Cys Glu Glu Trp Leu Glu Arg Glu Asp Ala Val Thr Tyr
99 100 105 110
101 Ser Arg Gly Phe Ser Lys Glu Pro Ile Phe Val Ser Gly Ala Asp Gln
102 115 120 125
104 Glu Trp Lys Ser Cys Ser Val Gly Cys Lys Phe Gly Phe Ser Gly Asp
105 130 135 140
107 Arg Lys Pro Asp Ala Ala Phe Gly Leu Pro Gln Pro Ser Gly Thr Ala
108 145 150 155 160
110 Ser Ile Leu Arg Ser Met Glu Ser Ala Glu Tyr Tyr Ala Glu Asn Asn
111 165 170 175
113 Ile Ala Met Ala Arg Arg Arg Gly Tyr Asn Ile Val Met Thr Thr Ser
114 180 185 190
116 Leu Ser Ser Asp Val Pro Val Gly Tyr Phe Ser Trp Ala Glu Tyr Asp
117 195 200 205
119 Met Met Ala Pro Val Gln Pro Lys Thr Glu Ala Ala Leu Ala Ala Ala
120 210 215 220
122 Phe Ile Ser Asn Cys Gly Ala Arg Asn Phe Arg Leu Gln Ala Leu Glu
123 225 230 235 240

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/913,858

DATE: 04/03/2002

TIME: 10:22:05

Input Set : A:\030560-057.ST25.txt

Output Set: N:\CRF3\04032002\I913858.raw

```

125 Ala Leu Glu Lys Ser Asn Ile Lys Ile Asp Ser Tyr Gly Gly Cys His
126           245           250           255
128 Arg Asn Arg Asp Gly Arg Val Asn Lys Val Glu Ala Leu Lys His Tyr
129           260           265           270
131 Lys Phe Ser Leu Ala Phe Glu Asn Ser Asn Glu Glu Asp Tyr Val Thr
132           275           280           285
134 Glu Lys Phe Phe Gln Ser Leu Val Ala Gly Thr Val Pro Val Val Val
135           290           295           300
137 Gly Ala Pro Asn Ile Gln Asp Phe Ala Pro Ser Pro Gly Ser Ile Leu
138 305           310           315           320
140 His Ile Lys Glu Ile Glu Asp Val Glu Ser Val Ala Lys Thr Met Arg
141           325           330           335
143 Tyr Leu Ala Glu Asn Pro Glu Ala Tyr Asn Gln Ser Leu Arg Trp Lys
144           340           345           350
146 Tyr Glu Gly Pro Ser Asp Ser Phe Lys Ala Leu Val Asp Met Ala Ala
147           355           360           365
149 Val His Ser Ser Cys Arg Leu Cys Ile His Leu Ala Thr Val Ser Arg
150           370           375           380
152 Glu Lys Glu Glu Asn Asn Pro Ser Leu Lys Arg Arg Pro Cys Lys Cys
153 385           390           395           400
155 Thr Arg Gly Pro Glu Thr Val Tyr His Ile Tyr Val Arg Glu Arg Gly
156           405           410           415
158 Arg Phe Glu Met Glu Ser Ile Tyr Leu Arg Ser Ser Asn Leu Thr Leu
159           420           425           430
161 Asn Ala Val Lys Ala Ala Val Val Leu Lys Phe Thr Ser Leu Asn Leu
162           435           440           445
164 Val Pro Val Trp Lys Thr Glu Arg Pro Glu Val Ile Arg Gly Gly Ser
165           450           455           460
167 Ala Leu Lys Leu Tyr Lys Ile Tyr Pro Ile Gly Leu Thr Gln Arg Gln
168 465           470           475           480
170 Ala Leu Tyr Thr Phe Ser Phe Lys Gly Asp Ala Asp Phe Arg Ser His
171           485           490           495
173 Leu Glu Asn Asn Pro Cys Ala Lys Phe Glu Val Ile Phe Val
174           500           505           510
177 <210> SEQ ID NO: 3
178 <211> LENGTH: 105
179 <212> TYPE: DNA
180 <213> ORGANISM: Artificial Sequence
182 <220> FEATURE:
183 <223> OTHER INFORMATION: Description of Artificial Sequence: cDNA
185 <400> SEQUENCE: 3
186 gaagccctga agcactacaa atttagctta gcgtttgaaa attcgaatga ggaagattat 60
187 gtaactgaaa aattcttcca atcccttggt gctggaactg tcctt 105
189 <210> SEQ ID NO: 4
190 <211> LENGTH: 35
191 <212> TYPE: PRT
192 <213> ORGANISM: Artificial Sequence
194 <220> FEATURE:
195 <223> OTHER INFORMATION: Description of Artificial Sequence: Mung bean

```

→ must explain genetic source
 - See error summary sheet
 item 11

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/913,858

DATE: 04/03/2002

TIME: 10:22:05

Input Set : A:\030560-057.ST25.txt

Output Set: N:\CRF3\04032002\I913858.raw

```

197 <400> SEQUENCE: 4
198 Glu Ala Leu Lys His Tyr Lys Phe Ser Leu Ala Phe Glu Asn Ser Asn
199   1           5           10           15
201 Glu Glu Asp Tyr Val Thr Glu Lys Phe Phe Gln Ser Leu Val Ala Gly
202           20           25           30
204 Thr Val Pro
205           35
208 <210> SEQ ID NO: 5
209 <211> LENGTH: 15
210 <212> TYPE: PRT
211 <213> ORGANISM: Artificial Sequence
213 <220> FEATURE:
214 <223> OTHER INFORMATION: Description of Artificial Sequence:n-terminal sequence
215   of tryptic peptide
217 <220> FEATURE:
218 <221> NAME/KEY: MISC_FEATURE
219 <222> LOCATION: (5)..(5)
220 <223> OTHER INFORMATION: Xaa = any amino acid
222 <400> SEQUENCE: 5
W--> 224 Lys Pro Asp Ala Xaa Phe Gly Leu Pro Gln Pro Ser Thr Ala Ser
225   1           5           10           15
230 <210> SEQ ID NO: 6
231 <211> LENGTH: 10
232 <212> TYPE: PRT
233 <213> ORGANISM: Artificial Sequence
235 <220> FEATURE:
236 <223> OTHER INFORMATION: Description of Artificial Sequence:n-terminal sequence
237   of tryptic peptide
239 <400> SEQUENCE: 6
240 Pro Glu Thr Val Tyr His Ile Tyr Val Arg
241   1           5           10
244 <210> SEQ ID NO: 7
245 <211> LENGTH: 13
246 <212> TYPE: PRT
247 <213> ORGANISM: Artificial Sequence
249 <220> FEATURE:
250 <223> OTHER INFORMATION: Description of Artificial Sequence:n-terminal sequence
251   of tryptic peptide
253 <400> SEQUENCE: 7
254 Met Glu Ser Ala Glu Tyr Tyr Ala Glu Asn Asn Ile Ala
255   1           5           10
258 <210> SEQ ID NO: 8
259 <211> LENGTH: 10
260 <212> TYPE: PRT
261 <213> ORGANISM: Artificial Sequence
263 <220> FEATURE:
264 <223> OTHER INFORMATION: Description of Artificial Sequence:n-terminal sequence
265   of tryptic peptide
267 <400> SEQUENCE: 8

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/913,858

DATE: 04/03/2002

TIME: 10:22:05

Input Set : A:\030560-057.ST25.txt

Output Set: N:\CRF3\04032002\I913858.raw

```

268 Gly Arg Phe Glu Met Glu Ser Ile Tyr Leu
269   1           5           10
271 <210> SEQ ID NO: 9
272 <211> LENGTH: 29
273 <212> TYPE: DNA
274 <213> ORGANISM: Artificial Sequence
276 <220> FEATURE:
277 <223> OTHER INFORMATION: Description of Artificial Sequence:primer
279 <220> FEATURE:
280 <221> NAME/KEY: misc_feature
281 <222> LOCATION: (3)..(15)
282 <223> OTHER INFORMATION: n = any nucleotide
284 <400> SEQUENCE: 9
W--> 285 gcngartayt aygcngaraa yaayathgc 29
288 <210> SEQ ID NO: 10
289 <211> LENGTH: 22
290 <212> TYPE: DNA
291 <213> ORGANISM: Artificial Sequence
293 <220> FEATURE:
294 <223> OTHER INFORMATION: Description of Artificial Sequence:primer
296 <220> FEATURE:
297 <221> NAME/KEY: misc_feature
298 <222> LOCATION: (14)..(17)
299 <223> OTHER INFORMATION: n = any nucleotide
301 <400> SEQUENCE: 10
W--> 302 crtadatrtg rtanacngty tc 22
305 <210> SEQ ID NO: 11
306 <211> LENGTH: 20
307 <212> TYPE: DNA
308 <213> ORGANISM: Artificial Sequence
310 <220> FEATURE:
311 <223> OTHER INFORMATION: Description of Artificial Sequence:primer
313 <220> FEATURE:
314 <221> NAME/KEY: misc_feature
315 <222> LOCATION: (6)..(6)
316 <223> OTHER INFORMATION: n = any nucleotide
318 <400> SEQUENCE: 11
W--> 319 tadatnswyt ccatytcraa 20
322 <210> SEQ ID NO: 12
323 <211> LENGTH: 20
324 <212> TYPE: DNA
325 <213> ORGANISM: Artificial Sequence
327 <220> FEATURE:
328 <223> OTHER INFORMATION: Description of Artificial Sequence:primer
330 <400> SEQUENCE: 12
331 ctggaactgt ccctgtggtt 20
333 <210> SEQ ID NO: 13
334 <211> LENGTH: 20
335 <212> TYPE: DNA

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/913,858

DATE: 04/03/2002

TIME: 10:22:06

Input Set : A:\030560-057.ST25.txt

Output Set: N:\CRF3\04032002\I913858.raw

L:224 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5
L:285 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9
L:302 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10
L:319 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11